DOCUMENT RESUME

ED 146 932 IR 005 403

AUTHOR Voges, Mickie; And Others
TITLE Recommended Systems for the Incremental Automation of the Morgue of "The Daily Texan."
PUB DATE Jun 76
NOTE 63p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Computer Programs; Indexes (Locaters); Information Retrieval; *Information Storage; *On Line Systems; *School Newspapers

ABSTRACT

A modular program is recommended for automation of the clippings file of "The Daily Texan" (student newspaper of the University of Texas at Austin). The proposed system will lead ultimately to on-line storage of the index, on-line storage of local staff-written news stories from the previous twenty-four months, micrographic storage for backup and retrospective search, and rapid retrieval techniques for accessing both the computer-stored articles and micrographically-stored articles. The system can be implemented on the EDP11/15 if the appropriate software and storage capabilities are available. It will provide several improvements to the present system including more complete indexing of news stories, more efficient document retrieval, more reliable file integrity, and easier access by the "Texan" staff to morgue files. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED146932

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RECOMMENDED SYSTEMS FOR THE INCREMENTAL AUTOMATION OF THE MORGUE OF THE DAILY TEXAN

Submitted to Texas Student Publications
(Mr. L. L. Edmonds, General Manager)

On the Initiative of the Study Group on
Automation of the Morgue of The Daily Texan

Graduate School of Library Science
University of Texas at Austin

June, 1976

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

E.B. Jackson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR 000 5 403

TABLE OF CONTENTS

Abstract iii

Foreword iv

Introduction 1

SUMMARY OF RECOMMENDATIONS 4
- System 1--Current System 5
- System 2--The Professionally-Staffed Morgue 5
- System 3--The Computer-Generated Printed Index 6
- System 3 Option--Automated Index 8
- System 4--Microfilm Jackets 9
- System 5--The On-Line Index 10
- System 6--On-Line Retrieval 11
- System 6 Option--Retrospective KWIC Index . 12

IMPLEMENTATION 13
- PERT Chart 14
- Flow Chart 16

APPENDIX A. THE MANUAL SYSTEMS 19
- System 1--Current System 19
- System 2--The Professionally-Staffed Morgue 20

APPENDIX B. THE MICROFORM AND AUTOMATED SYSTEMS 24
- System 3--The Computer-Generated Printed Index 24
- System 3 Option--Automated Index 28
- System 4--Microfilm Jackets 30
- System 5--The On-Line Index 31
- System 6--On-Line Retrieval 34
- System 6 Option--Retrospective KWIC Index . 35

APPENDIX C. SYSTEM COSTS 37
- 1--Current System 38
- 2--The Professionally-Staffed Morgue . . . 38
- 3--The Computer-Generated Printed Index . . 39
- 4--Microfilm Jackets and Computer-Generated Index 40
- 5--On-Line Index with Paper Clippings . . . 42
- 5a-On-Line Index with Microfilm Jackets . . 42
- 6--On-Line Index with On-Line Stories . . . 44
- Comparison of System Costs 45
- Summary of Costs on a Step-by-Step Basis . 46

APPENDIX D. REJECTED SYSTEMS/ALTERNATIVES 47
1'--Miracode 47
2'--Magnetic Tape Storage 48
3'--Computer Output Microform (COM) 49
4'--Questicon (ANCIRS) 49
5'--CDC 6600/6400 49

APPENDIX E. GLOSSARY 51

APPENDIX F. SAMPLES OF COMPUTER-GENERATED PRINTED INDEXES 52

Afterword 55

ABSTRACT

A modular program is recommended for automation of the clippings file of The Daily Texan. The proposed system will lead ultimately to on-line storage of the index, on-line storage of local, staff-written news stories from the previous twenty-four months, micrographic storage of stories for backup and retrospective search, and rapid retrieval techniques for accessing both the computer-stored articles and the micrographically-stored articles. The system can be implemented on the PDP11/15 if the appropriate software and storage capabilities are available. It will provide several improvements to the present system including more complete indexing of news stories, more efficient document retrieval, more reliable file integrity, and easier access by the Texan staff to morgue files.

FOREWORD

This study has been conducted and is presented by members of Professor E. B. Jackson's graduate seminar in computer-based information retrieval. In designing the structure of the course, Professor Jackson emphasized the practical problems of analyzing actual information systems. Consequently, with his advice, the class concluded that the best way to learn about such systems was to do a class project analyzing an existing system.

A number of study alternatives were presented by various class members. The proposal suggested by Tom Kirwin, a graduate student in journalism, was accepted by members of the class as the topic for study. Mr. Kirwin recommended a project to determine the feasibility of implementing an automated morgue for The Daily Texan. He believed Texas Student Publications would be highly receptive to a study of this nature.

During the course of the study, Mr. Kirwin served as chief liaison between the study group and Texas Student Publications. Members of the group interviewed various members of the staff of TSP, including Loyd Edmonds, Nick Cuccia, Art Renn, David Phillips, and Dorothy Krause. These individuals ably aided the research team in evaluating the state of the existing clippings file, the nature of the user and system needs, and the hardware and software requirements and capabilities of The Daily Texan's equipment. Their cooperation and encouragement are greatly appreciated by the study group.

INTRODUCTION

The purpose of the clippings file at The Daily Texan is to provide reporters with the accurate, concise information needed to write quality newspaper articles. This study group believes that modification and improvement of the current clippings file will significantly increase the reporters' ability to find pertinent information rapidly and efficiently. At the same time, the study group realizes that the production of a newspaper is a dynamic, daily process that cannot be curtailed while a new information storage and retrieval system is installed. Consequently, the system we propose is an evolutionary one which will begin building on the Texan's current information resources without denying the reporters those resources during the implementation period.

To accomplish this goal of increasing the benefits of information resources to the Texan staff with the least inconvenience, the following objectives were established:

1. The design of a simple and effective information storage and retrieval system for The Daily Texan.
2. The design of a system that allows modular growth from lower to higher levels of complexity with the least expenditure of TSP resources and with the least delay.
3. The design of a system compatible with the computing capabilities of The Daily Texan's PDP11/15 and not requiring support from outside (out-of-house) computing facilities.
4. The provision of an opportunity for the TSP staff and journalism students to familiarize themselves with information storage and retrieval techniques currently being used in modern newspaper libraries.

5. The design of a system enabling variation of the format of the file structure, whether on paper, microform or machine-readable form, so that user access to the files is not hindered during the implementation phase.

The study group feels that the implementation program proposed in this report meets the foregoing objectives and is consistent with the goal of an optimal newspaper morgue.

METHODOLOGY

The study group spent a worthwhile amount of time gaining a background knowledge of conventional and automated newspaper libraries. An extensive annotated bibliography on newspaper libraries was prepared by Mr. Kirwin and reviewed by the group. All group members toured the Austin American Statesman morgue and library and several members toured the libraries of the Dallas Morning News and Houston Chronicle.

At the inception of the study into the Texan's clippings library, the study group divided into four committees to look specifically at indexing, search, retrieval and cost. Committee assignments were made according to the academic and vocational strengths and experience of individual members. The committees served to investigate the requirements and capabilities of the various facets of a newspaper information system.

The bulk of information gathered by the committees came from personal interviews. Team members interviewed current users of the system, Texan management personnel, and computer technicians on the Texan staff. Various hardware vendors were called upon frequently to supply the group with technical specifications of both storage and retrieval devices. The University Computation Center often acted to verify information received from these sources.

At the conclusion of the committee inquiries, committee chairmen approached the group as a whole with sets of alter-

natives for improving the Texan's clippings morgue. The group, acting as a committee of the whole, weighed each alternative on its merits and shortcomings. Those alternatives which best met our requirements for an efficient and economic information system are incorporated into this final report.

This report then is a summation of the study group's findings and conclusions. It is organized into sections dealing with the current system, the proposed system, and the implementation of the proposed system with a tentative timetable. Finally, we have included a set of appendices detailing system descriptions, cost figures, and references. Organization in this manner should facilitate systematic appraisal of the research and the recommendations of the study group.

SUMMARY OF RECOMMENDATIONS

The study group has developed a sequence of recommended actions leading eventually to an automated morgue. These recommendations are presented in the form of morgue systems, each complete in itself:

System 1--The current system
System 2--The professionally-staffed morgue
System 3--The computer-generated printed index
    System 3 Option--Automated index
System 4--Microfilm jackets
System 5--The on-line index
System 6--On-line retrieval
    System 6 Option--Retrospective KWIC index

Each of these systems represents a significant level of improvement over the prior one; thus, System 2 is an improved version of the current system, System 1. In several cases, we have identified an optional step which may be appropriately implemented in conjunction with, or in place of, the recommended system. We have attempted to construct and order this sequence so that each step contributes to the final system. We have also attempted to eliminate those steps which would be superfluous to later systems.

In presenting these recommendations, we have discussed the resulting improvements in terms of the morgue itself. The morgue, of course, is merely a tool, a means to a better Texan. By providing easier access to a more complete store of information, an improved morgue contributes to more informative, thorough news reporting. As a dependable means to verify specific facts, names, dates, etc., the improved morgue also contributes to more accurate reporting.

Since each of the recommended systems is complete in itself, the developmental process can be concluded at any point. Although the study group has no vested interest in the development of any of these systems, we should hope that the Texan would give serious consideration to adopting the steps through System 4 at the minimum.

SYSTEM 1--THE CURRENT SYSTEM

The Daily Texan established a morgue in the fall of 1972. Currently there are 21,500 clippings in sixteen file drawers. The morgue's current growth rate is approximately 17,200 clippings a year, from 215 yearly issues; thus, storage space will soon become a major problem. Under the existing system, search and retrieval are difficult because the envelopes now in use contain either too many clippings (as a result of overly broad subject headings) or too few clippings (because many related terms are not cross-referenced). Other drawbacks of the present system include:

*No uniformity in establishing the envelopes' headings

*No consistency in alphabetizing and filing envelopes

*Lack of a card index to the morgue

*No control over use or misuse of the files

In an ideal manual system, empty envelopes or those seldom used would be purged, and the contents of large envelopes subdivided into more compact units. A card index and uniform set of cross-references would also facilitate retrieval. Finally, procedures would be instigated for orderly circulation of files.

SYSTEM 2--THE PROFESSIONALLY-STAFFED MORGUE

Results of a sample taken from six issues of the Texan, dated 27 February 1976 through 5 March 1976, showed that to

skim, select, clip and index news stories adequately requires two and one-half hours. In six days, 588 terms were selected, with an average of 2.58 terms per article. There is presently no standardized policy or practice of indexing; thus, ineffective subject headings, difficult search, and poor retrieval are inevitable. To improve the quality of indexing and thus the value and usefulness of the morgue to both staff and non-staff users, we suggest the following:

1. The creation of a half-time Library Assistant I position for a person with training in indexing or who has taken the Library School's course in indexing. This librarian would index each issue's articles and submit them to the secretary for clipping and filing.
2. The standardization of the clippings files, including adoption of uniform filing rules and reorganization of the current files in conformity with them.
3. The creation of an authority list for subject headings.
4. The compilation of a card index, leading from headings to files.
5. The creation of two procedural manuals: one to codify indexing procedures and one to govern the organization and circulation of files.

Implementation of these recommendations will result in more complete indexing, more compact and current files, improved search and retrieval capabilities, and easier training of new personnel. Should an automated system be chosen at a later date, these suggestions will facilitate the changeover to automation. (For further documentation of Systems 1 and 2, see Appendix A, "The Manual System.")

SYSTEM 3--THE COMPUTER-GENERATED PRINTED INDEX

The previous two systems benefit the Texan research staff only; this system represents the first level of improvement benefiting other researchers. A printed index can be

duplicated and furnished (or sold) to outsiders requiring retrospective access to Texan articles. The need for such a service is demonstrated by the Undergraduate Library's attempts to maintain a card index to the Texan. A newspaper morgue is an information resource whose value is underestimated. The newspaper is the best and often the only source of current information on many important subjects. Thus a newspaper has a responsibility to make its reference resources available to non-staff researchers. This is particularly true at an academic institution.

Unfortunately, meeting that responsibility can impose a real hardship on a newspaper morgue. The ANPA's Guidelines for Newspaper Libraries reports that the library staff may have to devote up to sixty-five per cent of its time to outside users. A printed index offers an ideal solution, for it can be duplicated at little cost and distributed to libraries and other organizations for the use of researchers. This allows the morgue staff to serve its primary users, the newspaper staff. Outsiders can simply be referred to the nearest library with a copy of the printed index.

A printed index could be produced using the Texan's present computer facilities. Multiple copies could be made serving as a search tool for non-staff researchers. The printed index would enable any user to locate quickly a specific article. The index would allow retrieval from clippings files, a later microform system or bound volumes since it would include both subject heading numbers and accession numbers with the date and page of publication of each article (for bound volumes or microfilm reels).

The index file would be maintained on an interchangeable disk cartridge compatible with the current disk drives. The file would be placed on-line each morning for sixty to ninety minutes for the librarian to index that day's articles. This should not interfere with production of the Texan. An alter-

native procedure would be to update the file once per week on a non-production day.

The index file would be used to generate automatically two versions of the index, a subject heading index and a complete index. The subject heading index would consist of an alphabetical listing of subject headings and the corresponding accession numbers for the clippings envelopes. Once the file was established, this listing would be produced only when a new subject heading was created or an old one was deleted. The complete index, a twelve-month cumulation of subject headings, headlines, and accession numbers for each article would be printed at the end of each semester. The December issue, indexing the current calendar year, would be permanently retained. Two copies of the subject heading index would be produced, one each for the TSP offices and the CMA building reporting/editing labs. Five copies of the complete index would be printed, two for the Texan and one each for the Undergraduate, Main and Communication Libraries.

In order to fulfill these goals, we recommend the following actions be taken:

1. Hire or contract a programmer for approximately six months.
2. Purchase additional disk cartridge for file storage and a magnetic tape for backup.
3. Enter at least twelve months, preferably twenty-four months, of subject headings from the current card index.
4. Eliminate card index as implementation of computer generated index is complete.

SYSTEM 3 OPTION--AUTOMATED INDEX

System 3 produces an index which must be searched manually. It also requires a librarian to update the index. Since automatically-produced keyword indexes rival manual indexing in facilitating retrieval, computer indexing becomes

a real alternative. It would thus be possible for the Texan to begin immediate development of an automatic indexing program which would extract index terms from the machine-readable text of articles.

A serious drawback to this approach, however, is that a keyword index cannot be readily searched manually. Index terms are single words, so a searcher would have to combine terms manually and locate common terms manually and then locate common entries in order to search for most subjects. The automatic indexing approaches which have so far been tested are suitable only for on-line searching, where the computer can perform the tedious task of combining terms. Therefore, implementation of an automated index should await the acquisition of the additional disk drive which would allow the file to be kept on-line for a sufficient length of time each day. Since the automated index cannot easily be used to generate a printed index, non-staff users would require access to a CRT (cathode ray tube) or teletype terminal. With no printed index for backup to the on-line system, the card index would have to be retained for that purpose.

Since several graduate students in the School of Communication are experimenting with automated indexing systems for newspapers, it is possible the Texan could obtain a workable program from this source at little or no cost. Since most of the programs developed for producing the printed index could be incorporated into the automated indexing system, we recommend the Texan delay a decision on this option and proceed instead with the development of the computer-generated printed index. This approach would allow comparison of the manual and automatic indexing systems before a final decision was made.

SYSTEM 4--MICROFILM JACKETS

Because of the constraints of physical space in the Texan morgue, it is inevitable that a change to some microform

system will eventually be necessary. Having examined several such systems, we believe that microfilm jackets provide a workable, efficient, and reasonably inexpensive microform system. We recommend the following:

1. The purchase of equipment necessary for maintaining a microfilm jacket system.
2. The use of a computer-generated printed index for access to the collection.
3. The gradual purging of the paper clippings files until only microfilm is used, both for current and retrospective research.

SYSTEM 5--THE ON-LINE INDEX

The Texan is currently considering the acquisition of additional disk storage capacity because of the heavy demands placed on current capacity by production requirements. Although the Texan is interested in acquiring more storage space than would be needed for the on-line index, the study group has designed a system which could be effected with just one 2.4 million character disk of the type currently used. We estimate that, by purging less important articles after twelve months, this disk memory would accommodate an on-line index to thirty-six months of those articles most likely to be required, or to two years of all indexed stories.

A relatively simple search program would suffice. The printed indexes would continue to be produced as a substitute during system down time and for retrieval of material older than three years. Since most news articles are written in the early afternoon and evening, we suggest the index be kept on-line from one until nine p.m. each day for searching. During other periods, the disk cartridge could be stored off-line so that the disk drive could be used for other purposes.

The following actions would be needed to implement this system:

1. Secure the services of a programmer, probably on a contracted basis.

2. Create a full-time position for a Library Assistant II.
3. Obtain an additional 2.4 million character disk drive and one additional interchangeable disk cartridge for backup.

SYSTEM 6--ON-LINE INDEX AND ON-LINE RETRIEVAL

Since most Texan copy is in machine-readable form at some stage in the production process, on-line retrieval of the complete text of articles is feasible if additional storage capacity is acquired. Furthermore, since the Texan is already considering the acquisition of additional capacity, and since the cost of disk storage is dropping, the Texan could conceivably realize such a capability within several years. Users would be able to retrieve text at a terminal and determine if a hard copy were needed. If a copy were desired, it could be printed out on the nearest line printer, presumably the TSP printer for the Texan staff. The printed index would serve as backup and for searching for material older than three years.

It is doubtful that the cost of the required disk storage could be justified simply as allowing on-line retrieval from a morgue text file. However, since the Texan may wish to expand computer capacity for other reasons, we would consider System 6 a realistic alternative. The text data base, it should be noted, would be available to other users for natural language processing research, including automated indexing experiments. With on-line storage costs dropping sharply, on-line retrieval of many types of documents soon could be a practical alternative. With the necessary storage capacity, the Texan could become center of research for such systems. This system would entail the following actions:

1. Obtain sufficient disk capacity for on-line storage of text.
2. Secure programming services for the necessary modifications to existing programs from previous systems.

3. Begin saving machine-readable text as soon as a decision is made to implement the system.
4. Eliminate clippings file, using microfilm reels or jackets for retrospective and backup search.

SYSTEM 6 OPTION--THE RETROSPECTIVE KWIC INDEX

The recommended systems do not overcome one major deficiency: no access to articles published prior to the implementation of the systems. One economical solution would be a KWIC (Key Word In Context) index to the headlines of past issues of the Texan. Although a KWIC index is only as good as the headlines, and usually identifies only a few terms per article, it is certainly an improvement over guessing at a date and then scanning all issues published about that time. It would be possible to produce a KWIC printed index to as many or as few past years of the Texan as desired. Such an index would be of some use to the Texan staff itself, and of even more use to non-staff researchers conducting historical research. Thus, some of the campus libraries might be willing to contribute to the funding of such a project, or the resulting index may be marketed for some return of cost.

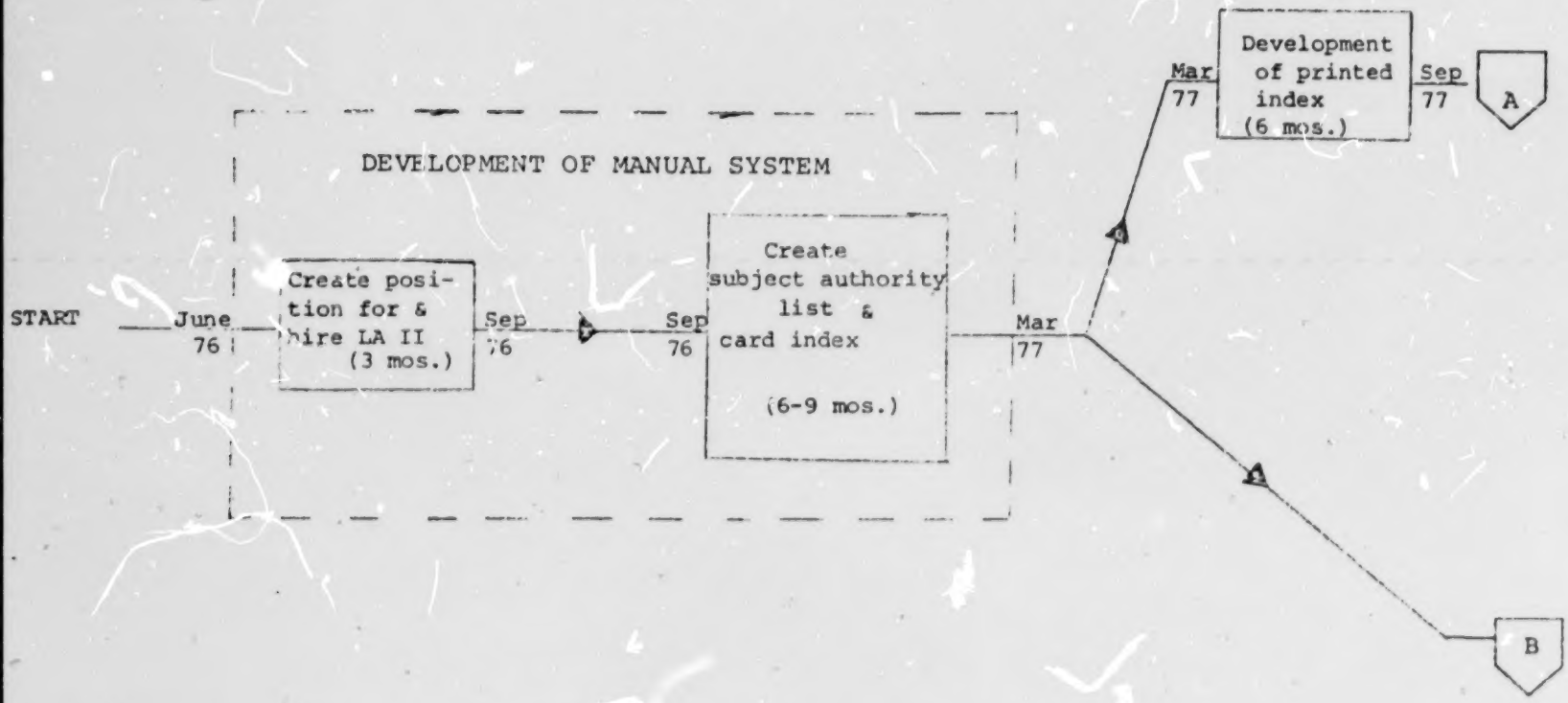
IMPLEMENTATION

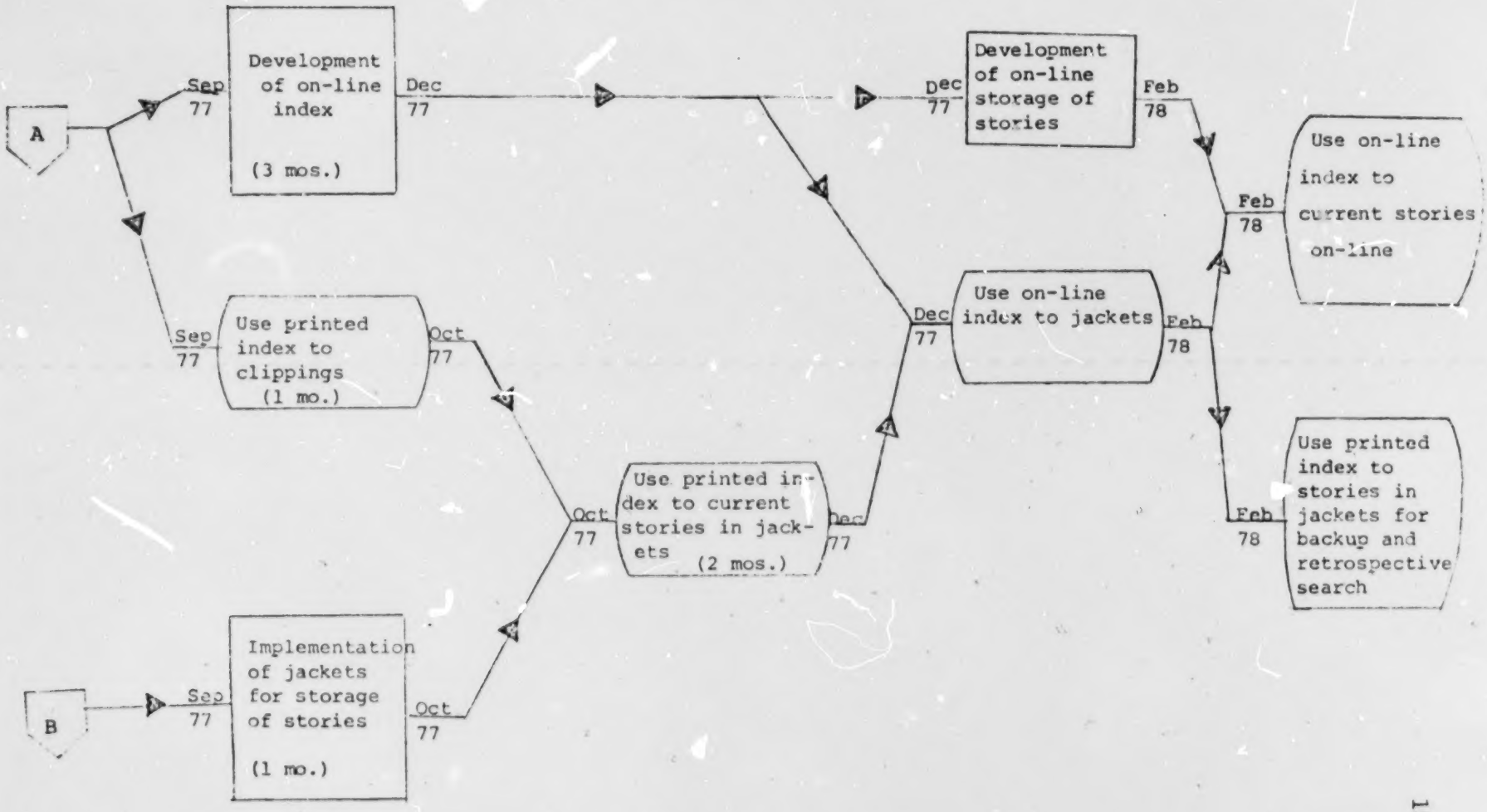
Implementation of the recommended systems is shown in the following two charts.

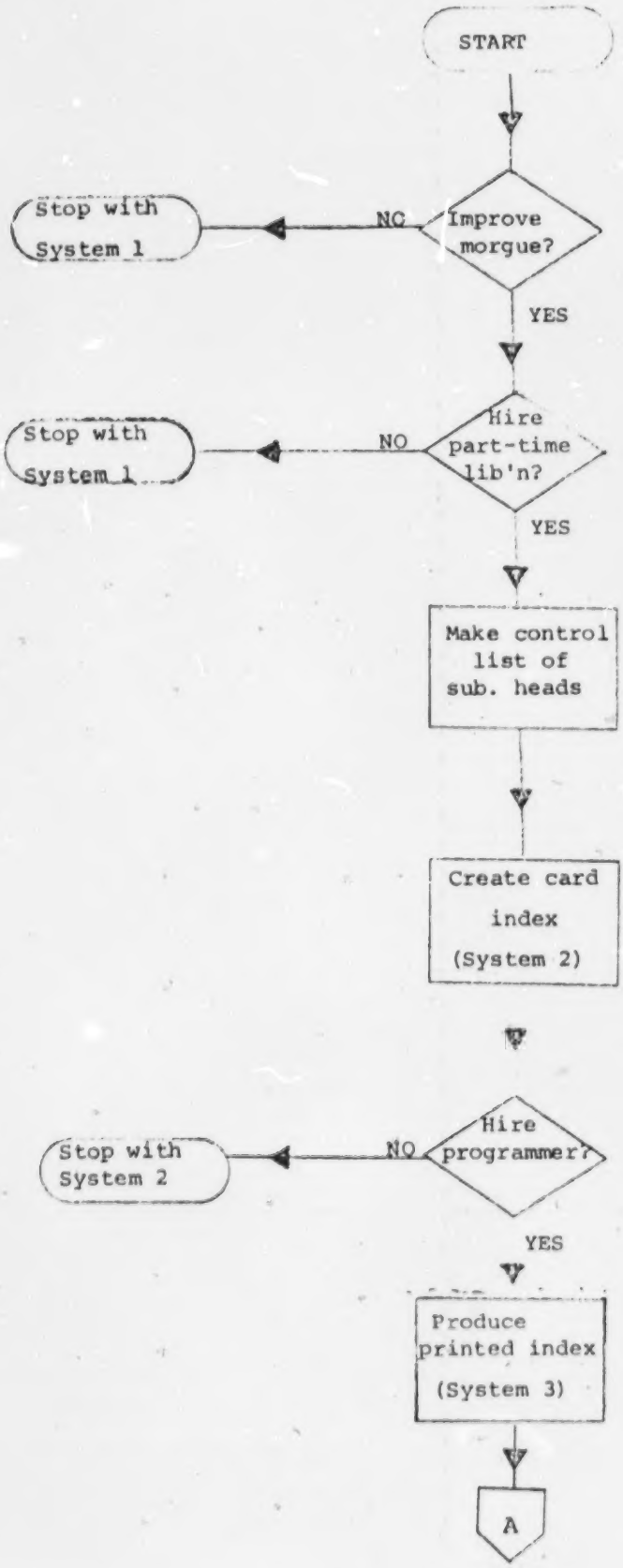
The PERT (Program Evaluation and Review Technique) Chart represents (in blocks) the phases of the implementation with arrows between the blocks indicating that a certain phase cannot be begun until some other phase is completed. The dates included are only rough approximations of what the study group estimates are feasible dates for the beginning and completion of the various phases. These dates may be changed if new situations arise which cause changes in the schedule. The only constraint in this respect is that if a date is changed, all dates coming after it must reflect that change. The PERT Chart follows the flow of the recommended systems without the options which are given.

The Flow Chart's main objective is to point out the various decision points which occur during the implementation of the systems. The advantages and disadvantages of each alternative that can be taken at each decision point are given in the explanatory appendices.

IMPLEMENTATION PERT CHART

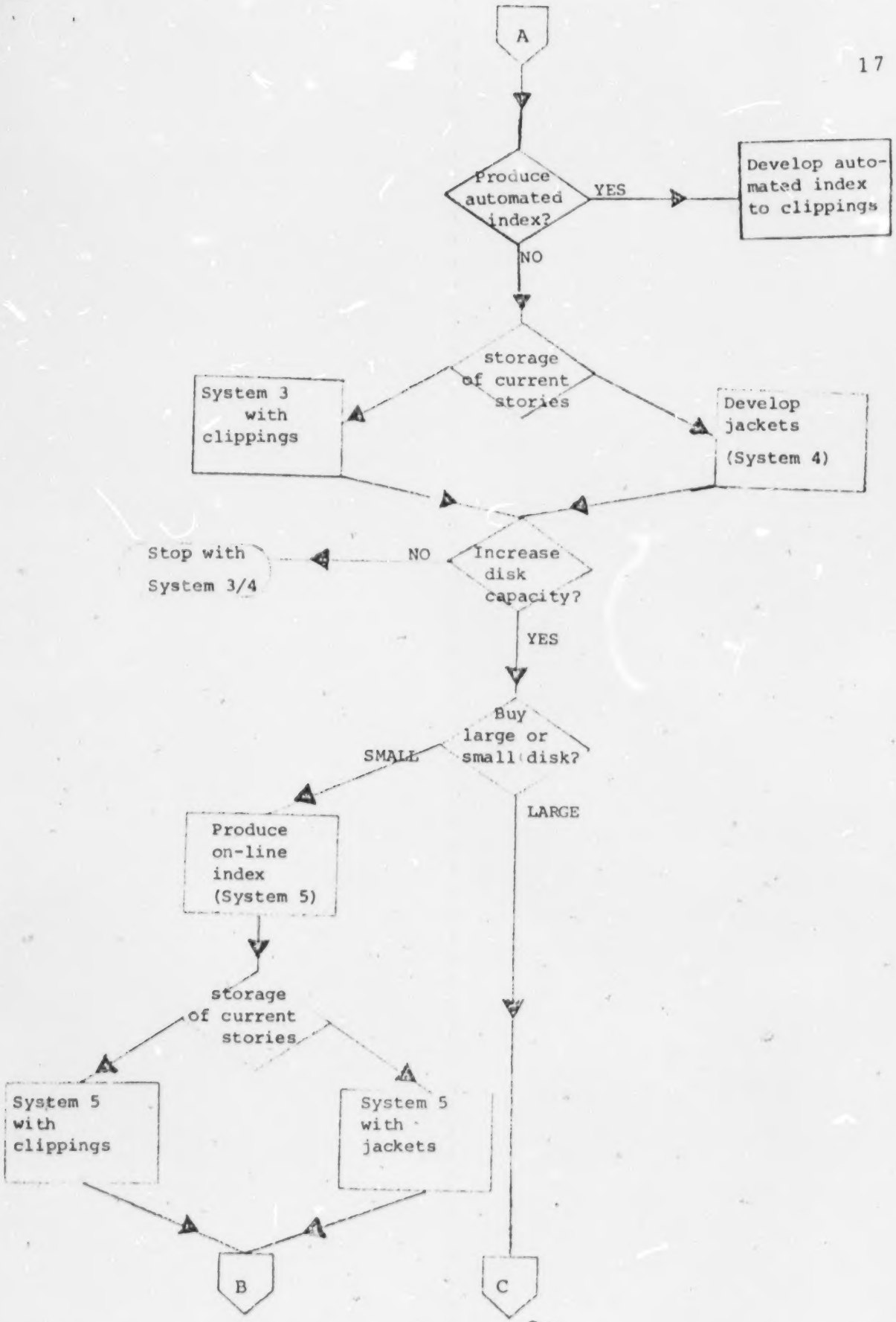


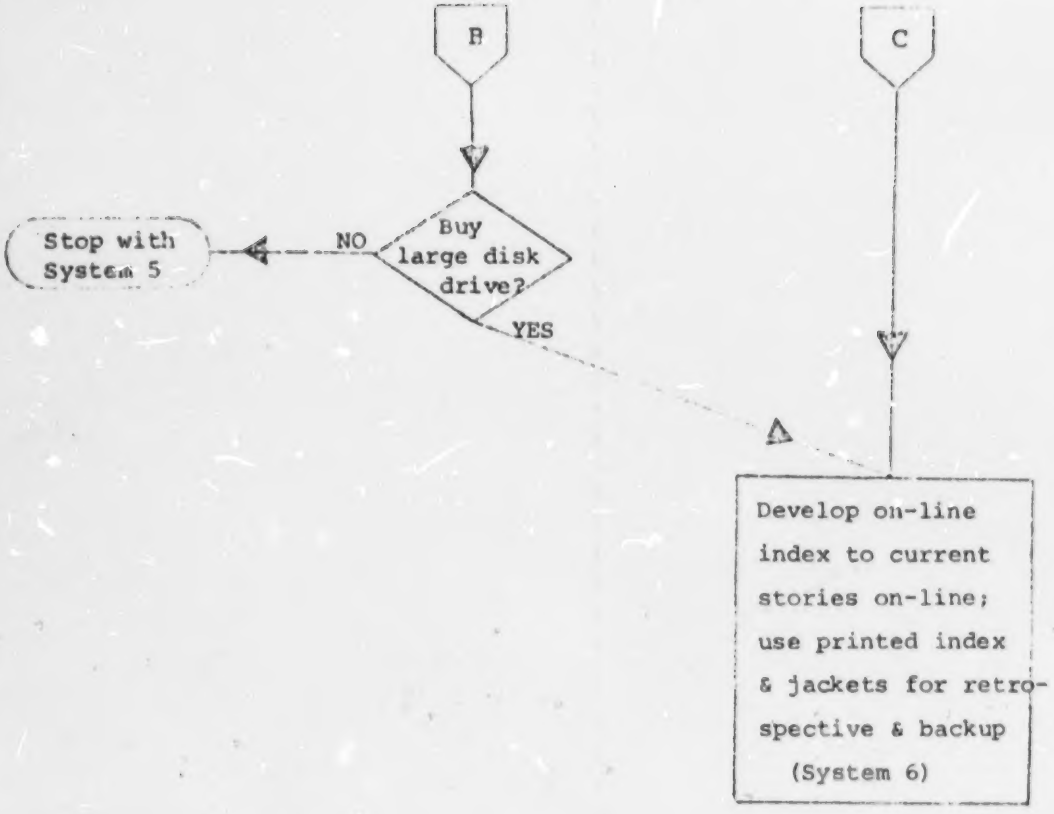

A
B
Sep 77
Development of on-line index
(3 mos.)
Dec 77
Sep 77
Use printed index to clippings
(1 mo.)
Oct 77
Sep 77
Implementation of jackets for storage of stories
(1 mo.)
Oct 77
Oct 77
Use printed index to current stories in jackets
(2 mos.)
Dec 77
Dec 77
Development of on-line storage of stories
Feb 78
Dec 77
Use on-line index to jackets
Feb 78
Feb 78
Use on-line index to current stories on-line
Feb 78
Use printed index to stories in jackets for backup and retrospective search



IMPLEMENTATION

FLOW CHART


A
Produce automated index?
YES
Develop automated index to clippings
NO
storage of current stories
System 3 with clippings
Develop jackets (System 4)
Stop with System 3/4
NO
Increase disk capacity?
YES
Buy large or small disk?
SMALL
LARGE
Produce on-line index (System 5)
storage of current stories
System 5 with clippings
System 5 with jackets
B
C


B
C
Stop with System 5
NO
Buy large disk drive?
YES
Develop on-line index to current stories on-line; use printed index & jackets for retrospective & backup (System 6)

APPENDIX A
THE MANUAL SYSTEMS

SYSTEM 1--THE CURRENT SYSTEM

The Daily Texan established a morgue (clippings file) in the fall of 1972. At present, the secretary of the Texan, Ms. Dorothy Krause, clips and files staff-written articles in envelopes labeled with general subject headings, personal names, place names, institution names, and similar indexing terms. Wire service material and syndicated columns are not clipped, although the morgue still houses a limited amount of such material from prior years. The morgue grows at the rate of eighty clippings a day, since an average of forty staff-prepared articles in each issue are clipped and filed under an average of two headings. At the present rate of growth, the morgue should increase by 17,200 clippings per year since approximately 215 issues of the Texan are published annually.

The current morgue, which covers a period of three and one-half years, contains an estimated 21,500 clippings in sixteen file drawers. Most of the file drawers in use are not full; it is estimated that the present morgue could be compressed into eight drawers. Each drawer can hold about 2500 to 3000 clippings. Even if all sixteen drawers continue to be available, space will soon be exhausted if the Texan begins to add 34,400 clippings a year to the morgue, since the file drawers can hold only an estimated total of 53,750 clippings.

Subject headings tend to be too broad. Some twenty-eight per cent of the envelopes with subject headings contained

fifteen or more clippings. The eleven envelopes in this category contained a total of 560 clippings, for a mean of 34.5 clippings per envelope. A user handed such a stack might have to invest too much time deciding which actually answered his query.

The user faces a different problem if he retrieves an envelope with only a few clippings. Since in the past it was the practice to file only one copy of each article, he could assume that other articles on the same subject might have been filed under a different heading, or perhaps a different version of the same heading ("Board of Regents" or "Regents, Board of"). A user would thus find it difficult to locate all the relevant clippings on such a topic.

In the ideal morgue, each article should be indexed under each subject which might be of interest to a user, including the names of the principle persons or institutions associated with the event reported. As a practical estimate, we suggest that articles should be indexed under an average of four headings, which might include several names and several general subject headings. From this estimate, we can project the growth rate of the ideal morgue for the Texan. Since the Texan now publishes about 215 issues per year, each containing about forty staff-written articles, 8600 articles would be indexed each year. If an average of four copies of each article are clipped, for filing under four headings, the morgue would increase by 160 clippings per day or 34,400 clippings per year.

SYSTEM 2--THE PROFESSIONALLY-STAFFED MORGUE

The conclusions to follow are the result of an unofficial test of indexing time and terminology. The test was run from 27 February 1976 through 5 March 1976 using the Texan issues of those dates. Indexing was done by one library school student without the aid of thesauri, authority lists or stop-

lists. Materials used included one copy of each day's paper, pens, and index cards.

Each day's paper was read and the suitability of each story for indexing was evaluated. Selected articles were then indexed by appropriate terms, which were recorded on index cards. Records were kept of number of articles indexed, time required to index, time required to list terms, total number of index terms assigned to each article, number of recurring terms within the issue and from day-to-day, and number of personal names used as terms. Also noted was the source of each story.

Results showed that to skim, select, and index the articles required two and one-half hours; listing terms required another two and one-half hours. For six days, a list of 588 terms was accumulated, averaging 2.58 index terms per article. Personal names were used as terms 247 of 588 times. Words in headlines often served as index terms; the average was four headline words per article.

The main conclusion to be drawn from this study is that one part-time employee can manually prepare and maintain an index to selected articles of The Daily Texan, given proper guidelines for the task. In the interest of the integrity of the clippings files and more complete indexing, it would be desirable to have the indexing done by a person with training in indexing methods.

In view of the above findings, the study group arrived at the following recommendations for improving the present morgue:

1. The Texan should create a half-time (20 hours per week) position for a qualified librarian. This individual would be responsible for indexing the articles in each issue of the Texan. The Texan's secretary would continue to clip and file the articles according to the librarian's instructions. Given the presence of the University's Graduate School of Library Science, it should be possible to keep this new position filled by students who have had prior experience or training in indexing

methods. Otherwise, the librarian could be required to take the Library School's course in indexing methods at the earliest opportunity.

2. The first task of the librarian should be to standardize the arrangement of the clippings files. This would entail adopting a set of filing rules, such as the American Library Association's filing rules, or the filing rules presently used by the Houston Chronicle. Next the librarian should review and reorganize the morgue, establishing cross-references, purging obsolete subject headings, and removing empty file folders. Maintenance of compact and current files is the objective of these activities.

3. The Texan should develop a codified list of subject headings (authority list). An authority list provides guidance to past indexing practices. It also enables the indexer to assign subject headings more consistently to a given article, thereby indexing more reliably articles on one subject under the same headings. In addition, the user of an authority list readily can understand the indexing scheme, and thus increase his chances of locating the required information. The present subject headings, combined with those developed for the Undergraduate Library's Texan index, are possible models; others include the Library of Congress headings, the New York Times terms, and Sear's List of Subject Headings.

4. A card index should be compiled, and users of the morgue should be required to search the card index rather than the clippings envelopes. There would be one card for each subject heading. Each card would be given a unique number which corresponded to an envelope with the same number. After using the needed clippings file, the user would return the envelope to be filed by the library staff. These steps will improve the accessibility of the files.

5. In order to standardize routines and facilitate training of personnel, two procedures manuals should be compiled. One manual should systematize the indexing process, including selection of the articles, assignment of subject terms, listing of terms in the card index and authority file, and preparation of file envelopes with matching

labels. The other manual should cover the organization and circulation of materials, including routines for filing, charging out, and refiling.

The above suggestions are interdependent and should be undertaken as a concerted effort rather than piecemeal. Implementation of the recommended modifications will benefit the Texan staff and other users, and the paper itself. We urge their adoption as soon as possible.

APPENDIX B
MICROFORM AND AUTOMATED SYSTEMS

This section presents a detailed description and rationale for the series of improvements to The Daily Texan news library or morgue that relate to the microform and/or automated systems previously proposed. The improved morgue systems described herein constitute a sequence of steps leading to the development of a totally automated information retrieval system for the Texan. Each system represents a significant improvement of at least one component--the indexing, searching or retrieving processes--of the prior system. Each system builds upon the improvements introduced in prior systems, but each is complete in itself. Thus, the developmental process could be ended at any stage.

SYSTEM 3--THE COMPUTER-GENERATED PRINTED INDEX

The improved manual system suffers from the major deficiency of providing no means to retrieve specific articles. A printed index which included the headline of each article indexed under a subject heading would correct this shortcoming. Since a printed index can be published in multiple copies, it can be distributed to all potential users: the Texan itself, the reporting/editing labs at the CMA building, the Undergraduate Library, Main Library, Communication Library, the Austin Public Library, perhaps even local newspapers, some campus staff agencies and some businesses.

The Daily Texan could use its current computer facilities for the production of a printed index. The computer programs

devised for the production of such an index would be retained for use in succeeding systems.

The Texan currently requires most of the available disk storage (5 million characters) for daily production. Thus, a computer-based index could not be kept on-line for a sufficient number of hours to allow on-line searching. In discussions with David Phillips, however, it was ascertained that sufficient on-line storage would be available in the morning hours for updating and printing an index.

The index files and computer programs would be stored off-line on magnetic tape, or on an interchangeable disk cartridge of 2.4 million characters, except when being updated. Each day of Texan publication, the file and programs would be placed on-line for sixty to ninety minutes to allow the librarian to update the index. For each article, the librarian would enter the headline and then the accession numbers of the appropriate subject headings. As the first task each day, this could be completed early enough in the morning to preclude interference with production of the Texan.

The index file would be used to produce two versions of the index. At the end of each semester, an index would be printed covering all articles published during the previous twelve months. This version of the index, called the complete index, would consist of an alphabetical arrangement of subject headings, each followed by the headline and accession number of each article indexed under that heading. The semester indexes would consist of approximately 350 pages of computer printout, with two columns of subject headings per page.

The accession number for each article would be the key to non-staff use of the index. The fact that the Undergraduate Library established a card index to the Texan provides evidence that there is demand by non-staff users for an index to the Texan. Although the secretary/librarian attempts to provide some assistance to outside researchers, she does not have time to do so on a large scale. Some risks are also incurred in opening

morgue files to other than staff members. An appropriately constructed accession number for each article can eliminate these problems. Each article would be assigned an accession number consisting of sections or fields for year, month, day, page number, column number. With this number and the headline, researchers can use the printed index to locate and retrieve any specific article from bound volumes or microfilm reels. The subject heading, plus the headline, would also enable the searcher to make a more accurate estimate of the pertinence of each article to his inquiry. This information will also serve the needs of the reporter or editor who must locate a specific article.

Furthermore, this information will be required for an on-line index (System 5). For that on-line index, the printed index will be needed as a substitute in case of computer or other equipment failure. In addition, outside users will still need the headline and accession number. Since this is an academic institution, it was felt that the Texan owes some obligation to such users. For staff use, a second version of the printed index will be produced as often as needed. With this printed index, the Texan will still be using clippings files for current retrieval. Staff will simply need to look at a list of all current subject headings, plus the subject heading number. This subject heading index would cover the last three years of articles (or whatever period was adopted as the limit for current retrieval). It would be printed, in two copies, whenever a new subject heading or subheading was created, about twice a month. Only staff-written articles would be indexed. At the current rate, there would be 8600 articles to index per year. At an average of four subject headings per article, an estimated 34,400 entries would be added each year.

General indexing theory suggests an average of six entries per subject heading. For newspapers, however, it was felt that this figure was too low. The typical index provides

bibliographic references, which will often require the searcher to examine a different issue of a journal for each entry. Newspaper retrieval is much faster, for the searcher retrieves a file containing all articles indexed under a given subject heading. An additional consideration is the fact that most articles are relatively short, only three to four hundred words. The study group therefore adopted the average of fifteen entries per subject heading, which would require the use of about 2500 subject headings per year, or a total of 7500 subject headings and subheadings for the three-year index.

The complete index file would consist of three years of subject headings, each with a clippings file accession number of four digits, and one year of headlines (up to 50 characters) with an accession number for each. Thus the storage requirements would be as follows:

7500 subject headings at 20 characters each	150,000	characters
7500 subject accession numbers at 4 digits each	30,000	"
8600 headlines at 50 characters each	430,000	"
8600 accession numbers at 11 characters each x 4 per article	554,400	"
TOTAL	1,164,400	"

Using a 2.4 million character disk pack, this would leave about one million characters of storage for the necessary computer programs.

According to David Phillips, implementation of the printed index would require the services of a full-time programmer for about six months. If System 5 and/or System 6 is also to be developed, the requirement of additional programming would argue for creation of a position for one year, beginning in September, 1977. Once the programs were completed, the appointment could probably be suspended since regular programming staff could maintain the programs.

The required programs would be modular, so that those used to handle updating of the printed index could later be used for updating the on-line index. The programs should also be designed in anticipation of the additional storage requirements for the on-line index file. In this way, all of the programs developed to implement this system will be retained for use in Systems 5 and 6.

SYSTEM 3 OPTION--AUTOMATED INDEX

The feasibility of automatic indexing of machine-readable text has been demonstrated in a variety of experimental studies within the last ten years. Such indexes, called keyword indexes, rest on computer extraction of index terms (single words) from the text of articles. Comparison of retrieval performances from the same set of documents using both a keyword index and a manually-produced index, including one study of news material, have repeatedly demonstrated that there is no inherent advantage in manual indexing: retrieval performance is essentially the same no matter which index is used for searching. Perhaps the best evidence for the quality of computer indexing is the fact that at least one computer manufacturer is now marketing such a computer program. This program, however, requires the use of an IBM System/370 with a minimum of 240,000 characters of core storage. At some time in the near future such a program may become commercially available for the PDP 11, or for one of the other University computers. In the latter case, more consideration should be given to creating the on-line index on an outside computer facility.

A serious shortcoming of a keyword index, however, is the fact that, for all practical purposes, it must be searched on-line. Since each index term is a single word, a print-out of such an index would be monstrous. Furthermore, manual searching of the index would be time-consuming since the user would have

to coordinate terms. Many index terms would occur so frequently in the document collection that there could be hundreds of entries for any single term. The user would have to locate a second term in order to combine two that identified the specific topic he wished to retrieve. He would then have to compare the accession numbers under the two terms and select the common ones. If he still have too many articles, he would have to find a third term and repeat the procedure.

If such an on-line index were adopted, a manual backup would not be available, unless the librarian continued to maintain a card index. This would certainly appear to be wasteful.

The fact that several graduate students in the School of Communication are experimenting with automatic or computer-assisted indexing of newspaper articles deserves special consideration. The Texan may be able to acquire from one of these students a program for automatic indexing at little or no cost. At least one of these experimental programs is designed to extract phrases, an approach which would be more suitable for production of a backup printed index.

In the final analysis, the only drawbacks to implementation of this option would be the lack of a printed index as a measure of security and the fact that only people with access to the University computing system would be able to use the on-line index so produced. Furthermore, it would appear that the Texan would have to hire a programmer to design such an index, so there would be little, if any, cost advantage in going to the computer-produced index in preference to the on-line index produced by a librarian at this time. Other such programs may become commercially available in the near future, so the wisest course would probably be to proceed with the steps outlined for System 5. The programs so produced could be used, with a Boolean search capability added, for searching either type of on-line index. The programs for updating the index file could also be retained, for they would not be affected by the method of index term selection.

SYSTEM 4--MICROFILM JACKETS

The microfilm jacket system is essentially a microfilm version of a clippings file. A plastic jacket or microfilm holder is used to store individual frames of microfilm. One copy is filmed of an article for each subject heading assigned to it (this has been estimated a four per article). After the film is processed, the copies of the article are inserted into the appropriate microfilm jackets corresponding to the subject headlings by which the article had been indexed (a jacket loader is used for this process). A user of the system first searches the index to locate the number of the appropriate subject (this corresponds to a jacket). The jacket is placed in a microfilm reader/printer, and articles are examined in much the same way that one would look through a stack of clippings contained in an envelope. If a hard copy is required, the print button on the reader is pressed.

To be of value, a morgue must be current. We therefore recommend that equipment be purchased to allow the librarian to expose the microfilm. Stories, having been indexed and clipped would be photographed by the librarian. One one copy of the clipping would be clipped because it could be photographed as many times as was necessary to meet the requirements of the indexing. (If a story were indexed under four subject headings, it would simply be photographed four times.) Exposed film would be processed locally (out-of-house), thus avoiding the necessity of purchasing extra equipment and chemicals and of hiring or training special personnel. It would be possible to have jackets current to within a week or two.

The purchase of two reader/printers and a storage cabinet would also be necessary for the adoption of the jacket system. Additionally, we recommend purchase of a film copier in order that users may make inexpensive film copies of the jackets that may be used on readers at other locations and that the librarian

may be able to produce duplicate copies of a file if the need arises.

The clippings file and the jacket system would need to be maintained side-by-side for some time as the change was being made. Eventually, through purging the clippings files, the jacket system would completely replace the same.

Because of the simple requirements of physical space of a clippings file, we believe that some microform system is inevitable for the morgue of the Texan. Of systems currently available, the jacket system is the most appropriate for the Texan morgue for several reasons:

1. It is relatively inexpensive.
2. Jackets are very similar to the clippings files already in use and familiar to the users; the change would be easy for them to make.
3. This system does not require any great amount of expertise to work--extra personnel would not be required.
4. Because jackets are collections of clippings all pertaining to a specific subject (within a particular jacket) there is a browse capability which is beneficial to researchers.
5. The jacket system is quite flexible in that it is relatively easy at any time to change a subject heading, split a file which is too large into more specific subject headings, add new subject headings and delete old or inappropriate subject heading.

SYSTEM 5--THE ON-LINE INDEX

The Texan's computer system currently includes thirteen cathode ray tube (CRT) terminals, eight located in TSP and five in the reporting/editing labs of CMA. An on-line index would thus be accessible to reporters or editors in either location. When the Texan computer is added to the University computing system, other researchers would also gain access to the on-line data base.

According to David Phillips, a programmer would need about three months to create the additional computer program for on-line searching of the index. If Boolean logic were incorporated, six months would probably be necessary for program implementation. The programs developed for System 3 would be retained and used to produce the subject heading index as support of the on-line system for staff members. The complete index would still be produced on the same schedule to serve the needs of users having no access to the on-line index. The annual cumulations would be kept by the Texan for retrospective retrieval (articles more than three years old). The on-line index would be compatible with retrieval of articles from clippings or microfilm jackets by staff members or from bound volumes or microfilm reels by non-staffers.

Ideally, the on-line index would be expanded to encompass a three-year period. Most users would then have no need for the printed index. An index of this size would require an additional 1.5 million characters of disk storage over and above that required for System 3, or a grand total of 2.6 million characters of storage. While references to all articles would be retained for at least twelve months, in order to prepare the complete printed index, the librarian could begin to purge items from the index after this period. The computer could be used very readily to keep track of which subject headings were frequently searched, and this information could be used by the librarian in removing entries from those subject headings seldom or never used.

An additional disk drive would be required to maintain an on-line index. David Phillips believes the TSP computer disk storage capacity is near the saturation point. He has been investigating the possibility of obtaining additional disk storage to meet growing Texan production needs and to make the TSP system more accessible by researchers, one of the purposes for which it was obtained. Even if it were necessary

to purchase the additional 2.4 million character disk storage primarily to support the on-line index, we estimate that the index would need to be on-line about eight hours per day. By using interchangeable disk packs, the additional disk storage capacity could be used for other purposes simply by removing the disk pack used for the index files.

A relatively simple search program would probably be sufficient for most needs. An index with subject headings assigned by a qualified librarian could dispense with a Boolean search capability. The index would allow a user to search in two ways. The individual could browse, simply selecting a letter of the alphabet and then scanning all subject headings beginning with that letter. Or, the individual could enter what he knew or thought to be a subject heading and a match would be made against the index file. Once the individual had located what he believed to be the appropriate heading, he could request a list of headlines to verify his selection or to obtain the accession number of one or more articles. If the individual were a staff member, he would simply note the clipping or jacket file number for the subject heading, and then retrieve the appropriate jacket or envelope.

Since the computer programmer would need to work closely with the librarian in the development of the on-line systems, we recommend that a full-time position be created if System 5 is implemented. Since equipment acquisition may not be possible until the 1977-78 fiscal year, and since it will take some time to implement the previous system, this full-time position would probably not be required until September 1977. The full-time librarian (Library Assistant II) would take over the responsibility of conducting demonstrations and instructing the staff in the use of the on-line index.

SYSTEM 6--ON-LINE RETRIEVAL

Since the Texan now prepares all copy in machine-readable form for typesetting, even though some of this material does not pass through the TSP computer, a completely automated retrieval system is feasible for recent material (articles less than three years old). A major increase in storage capacity would be required to implement such a system, which could probably not be justified simply to automate the morgue. Since the TSP computer is also dedicated to support research, and since the required capacity for on-line retrieval of text could also be used for other purposes, the study group considered this a legitimate alternative.

The Texan should create a method of on-line retrieval which would enable morgue users to examine the complete text of articles at CRT terminals. In a very short time, a user would thus be able to make a final accurate judgment of the importance of a particular article. At his command, important articles could then be copied by the line printer.

Such a facility would require the purchase of a disk drive capable of storing up to sixty million characters. This would accommodate storage of all text for two years or, according to computer-generated statistics, the retention of one complete year (the most recent) and two years of those articles most likely to be retrieved. Thus, the on-line text file would contain the complete text of every article identified in the on-line index. According to the DEC representative in Austin, an 88-million character disk drive is now available for the PDP 11. This drive would certainly offer the necessary storage space for the on-line text file, leaving substantial machine memory for other uses. In addition, since interchangeable disk packs are available, the disk pack consisting of the index, text and programs for the morgue could simply be removed during the hours the morgue is not being used on-line. Thus, the entire capacity of the drive would be available during those hours for other purposes.

produce a printed KWIC (Key-Word-In-Context) index to past issues of the Texan. This index could be issued in annual cumulations for as many past years as desired. Looking at the processing of a file of titles about one-fourth the estimated size of one year of Texan headlines (8600 headlines at fifty characters each), a rough estimate of costs of computer time and supplies for an annual KWIC index (on the DEC 10) would be $100. Keypunching costs would be about $90; therefore, the Texan could issue these retrospective KWIC indexes for around $200 for each annual cumulation. Such an index would not only be useful to the Texan itself, but would be of even more use to non-staff personnel doing historical research. Thus, some of the campus libraries might be willing to contribute to the funding of such a project. An alternative would be to purchase a KWIC program for the PDP 11/15, or have one written for in-house production.

This system demands a standard procedure for recording any changes/corrections to text made manually so that the librarian could enter the corrections in the machine-readable text file.

The computer-printed index would continue to be used as backup and for outside users.

The clippings files would be eliminated. Microfilm storage would be used for older material and for backup to the on-line system. The printed index, of course, would make this material accessible.

With careful planning, it should be possible to initiate on-line retrieval as soon as the programs are developed. The text file could be established on tape almost immediately by copying each day's articles from the disk each evening before purging the disk files. Allowing space for end-of-file marks, we estimate that a twenty-four-foot tape can carry about thirty-six million characters. Since a year of the Texan is estimated to consist of twenty-nine million characters of text and headlines, a text file of fifteen months could fill a tape. This file would be useful to researchers now, and would be ready to become part of the on-line retrieval system as soon as the programs and disk capacity are available. Otherwise, the text file would have to be established after the system was developed and would be of little use during the first year.

SYSTEM 6 OPTION--THE RETROSPECTIVE KWIC INDEX

Until System 3 is implemented, there will be no means of conducting an efficient search for material about an older subject. The only recourse could be to scan each issue about the time of publication, assuming one knows the approximate date. If a searcher is looking for general background material, he has a more difficult task since he will often not have a date with which to work. An economical solution would be to

APPENDIX C
SYSTEM COSTS

Each system presented as an alternative in the body of the study is shown here in terms of start-up costs and annual operating costs. A significant cost that can only be determined at the time of implementation is the cost of a software program. Costs of these programs vary radically depending on the requirements of the program and where the programming is done: in-house or out-of-house. Systems requiring such programming are designated by an asterisk beside the title.

The following need to be noted in considering all cost figures:

1. No account of inflation is taken.
2. Depreciation is not considered in yearly operating figures.
3. Physical plant and office supplies costs are not shown.
4. When necessary, costs are rounded off.
5. All clerical costs refer to the University of of Texas classified 1975/76 schedule. Fringe benefits, such as sick leave, vacation or insurance, are not included.
6. Where computer costs are part of present overhead, they are not specified.

Cost figures are shown for the following systems:

1. Current system
2. The professionally-staffed morgue
3. The computer-generated printed index
4. Microfilm jackets and computer-generated index
5. On-line index with paper clippings

5a. On-line index with microfilm jackets

6. On-line index with on-line stories

1. Current System

Articles are clipped and filed in envelopes. No index to the files is provided.

1. Start-up costs		none
2. Operating costs for one year		
A. Staff labor		
1 1/2 hours clipping/filing a day at 20% of annual salary ($6,672)	$ 1,334	
B. Equipment operating costs	none	
	$ 1,334	total

2. The professionally-staffed morgue

Articles are clipped and filed in envelopes. A part-time Library Assistant I indexes the articles and establishes the authority lists. A manual card index provides access to the clippings.

1. Start-up costs		none
2. Operating costs for one year		
A. Staff labor		
1. 1 1/2 hours clipping/filing a day at 20% of annual salary ($6,672)	$ 1,334	
2. 4 hours indexing a day at 50% annual salary ($6,672)	3,336	
B. Equipment operating costs	none	
	$ 4,670	total

3. The computer-generated printed index *

Articles are clipped and stored in envelopes. A part-time Library Assistant I does the indexing. A printed index provides access to the envelopes.

The index is arranged alphabetically by subject. Pertinent articles under each subject heading are listed by an accession number and the headline of the article. The eleven-digit accession number indicates the date, page and column of the article so that the printed index may also be used to enter bound volumes of The Daily Texan or any chronologically microfilmed issues.

1. Background computations
   - A. Number of characters in a printed index to one year's stories
     1. 8600 stories a year indexed under 4 subject headings each = 34,400 entries
     2. 34,400 entries x 11 characters each = 378,400 characters of entries
     3. 34,400 entries ÷ 14 entries per subject heading = 2500 subject headings
     4. 2500 subject headings x 20 characters each = 50,000 characters of subject headings
     5. Annotations to subject headings are 50 characters x 8600 titles = 430,000 characters
     6. 2500 subject headings x 4 characters of subject accession numbers = 10,000 characters
     7. 378,400 characters of entries + 50,000 characters of subject headings + 430,000 characters of annotations = 858,000 characters a year in index
   - B. Number of pages in a printed index
     1. One line holds 136 characters
     2. Subject headings are 20 characters long. Each entry consists of an 11-digit accession number and a 50-character annotation. Each subject heading and subsequent entries will use less than 1/2 line.

3. 2500 subject headings x 14 entries each at 1/2 line each (2 columns down each page) = 17,500 lines

4. To allow margins and to avoid splitting entries, 50 lines per page are allowed. 17,500 lines ÷ 50 = 350 pages

2. Start-up costs

Purchase 7.4 million character interchangeable disk pack	$ 100	
	$ 100	total

3. Operating costs per year

A. Staff labor

1. 1 1/2 hours clipping/filing a day at 20% of annual salary ($6,672)	$ 1,334	
2. 4 hours indexing a day at 50% annual salary ($6,672)	3,336	
3. 145 hours a year to keyboard index at $3.21 per hour = $465 (858,000 characters of index - 100 characters keyboarded per minute)	inc. in #2	

B. Cost of printing the index

350 pages x $ .007 per page = $2.45 x 5 copies x 3 printings a year	37	
	$ 4,707	

* Software program required--see introductory section of this appendix.

4. Microfilm jackets and computer-generated index *

Articles are clipped and filmed on the jacket microfilm by subject heading. Indexing is done by a part-time Library Assistant I. Two printed indexes are provided. The jacket index gives each subject heading and its corresponding jacket number. The complete index gives each subject heading, its

jacket number, and all entries by chronological accession number and the story's headline (50-character annotation).

1. Start-up costs

1 camera	$ 900	
1 film copier	3,000	
1 jacket loader	1,800	
2 reader/printers	1,990	
1 cabinet	400	
	$ 8,090	total

2. Operating costs for one year

A. Staff labor

1. 1 1/2 hours clipping articles and filing jackets at 20% annual salary ($6,672)	$ 1,334	
2. 4 hours of indexing at 50% annual salary ($6,672)	3,336	
3. 2 hours to expose reel of film x 17 reels a year x $3.21 per hour.	110	
4. 145 hours a year to keyboard index at $3.21 per hour = $465 858,000 characters of index - 100 characters keyboarded a minute	inc. in #2	

B. Equipment operating costs

1. 17 reels per year at $2 each for developing	$ 34	
2. 17 reels per year at $9 each	153	
3. 2500 paper prints a year at $ .05 each	125	
4. Cost of printing the indexes 150 page jacket index x 2 copies x 21 printings a year x $.007 per page	44	
350 page complete index x 5 copies x 3 printings a year x $ .007 a page	37	
	$ 5,173	total

* Software program required--see introductory section of this appendix

5. On-line index with paper clippings *

Articles are clipped and filed in envelopes (the microfilm jacket system not having been implemented). The daily preparation and maintenance of the on-line index is done by a full-time Library Assistant II. Access to the stories is gained by an index of subject headings with eleven-digit chronological accession numbers and annotative headlines.

1. Start-up costs

A. 1 disk cartridge stores 2,500,000; a 3-year index is 2,398,000 characters; 1 disk is $100 plus a backup disk = $200	$ 200	
B. Disk drive = $5,000 the prorated cost of the morgue's use of the disk is 8 hours a day = 1/3 of $5,000 = $1600	5,000	
	$ 5,200	total

2. Operating costs per year

A. Staff labor		
Full-time Library Assistant II	$ 8,436	
B. Equipment operating costs--negligible		
C. Printed index		
150 page subject number index x 2 copies x 21 printings a year x $.007 per page	44	
350 page complete index x 5 copies x 3 printings a year x $ .007 a page	37	
	$ 8,517	total

* Software program required--see introductory section of this appendix

5a. On-line index with microfilm jackets *

Articles are clipped and filmed on jacket microfilm. The indexing and handling of the on-line index is by a full-time Library Assistant II. The on-line index is the complete

index. A backup printed jacket index is printed three times a week and a backup complete index is printed every three months.

1. Start-up costs

A. Jackets		
1 camera	$ 900	
1 film copier	3,000	
1 jacket loader	1,800	
2 reader/printers	1,990	
1 cabinet	400	
B. Disk		
1 disk cartridge + 1 backup disk cartridge	200	
1 disk drive	5,000	
	$13,290	total

2. Operating costs for one year

A. Staff labor		
Full-time Library Assistant II	$ 8,436	
B. Equipment operating costs		
1. 17 reels film per year at $9 each	153	
2. 2500 paper prints a year at $ .05 each	125	
3. 17 reels per year at $2 each for developing	34	
4. Cost of printing indexes 150 page jacket index x 2 copies x 21 printings a year x $.007 per page	44	
350 page complete index x 5 copies x 3 printings a year x $ .007 a page	37	
	$ 8,829	total

* Software program required--see introductory section of this appendix

6. On-line index with on-line stories *

Stories are indexed by a full-time Library Assistant II. Two years of stories are on-line. Stories are also stored (archives) on microfilm. Access to these stories is through the on-line index and/or the printed indexes.

1. Computation of number of characters in stories

   130,000 characters of stories and headlines per issue

   130,000 characters x 215 issues per year = 28,868,000 characters a year

   28,868,000 x 2 years = 57,736,000 characters

2. Start-up costs

   Purchase of a 88 million character disk drive and control unit = $35,000. The prorated cost of the morgue's use of the disk is:
   --58,604,000 characters is 67% of unit
   --67% of $35,000 = $24,450
   --disk is used 8 hours a day
   1/3 of $24,450 = $7,816 $35,000

   $35,000 total

3. Operating costs for one year

   A. Staff labor
      Full-time Library Assistant II $ 8,436

   B. Equipment operating costs--negligible

   $ 8,436 total

4. Backup systems

   A. Printed indexes $ 81

   B. Microfilm jacket storage
      Film 153
      Prints 125
      Developing 34

   $ 393 total

* Software program required--see introductory section of this appendix

COMPARISON OF SYSTEM COSTS

System	Start-up Costs	Operating Costs (1 year)
1. Current System	$ none	$ 1,334
2. Professionally-staffed morgue	$ none	$ 4,670
3. Computer-generated printed index	$ 100	$ 4,707
4. Microfilm jackets and computer-generated index	$ 8,090	$ 5,173
5. On-line index with paper clippings	$ 5,200	$ 8,517
5a. On-line index with microfilm jackets	$13,290	$ 8,829
6. On-line index, on-line stories --plus-- backup systems and index	$35,000	$ 8,436 393 $ 8,829

italics indicate option

SUMMARY OF COSTS ON A STEP-BY-STEP BASIS

Current system			
1. Clipping/filing			$ 1,334
Professionally-staffed morgue			
1. Add half-time employee	+$ 3,336		3,336
Addition of printed index			
1. Cost of printing index	+	37	
2. Purchase of disk	+	100	137
Addition of jacket storage			
1. Cost of equipment	+	8,090	
2. Film, prints & developing	+	312	8,402
Put index on-line			
1. Disk purchase	+	5,200	
2. Eliminate half-time employee	-	3,336	
3. Add full-time employee	+	8,436	10,300
Put stories on-line			
1. 88 million character disk and control	+ 35,000		35,000
			$58,509

APPENDIX D
REJECTED SYSTEMS/ALTERNATIVES

1'. Miracode; Eastman Kodak--Tony Byers

The document (clipping) is placed on the camera screen, an accession number is punched into the alphanumeric keyboard (encoder), and the film recording is made. A reel may be partially or completely developed in fifteen minutes. There is no real need to clip articles as whole pages could be filmed at once. The accession numbers could indicate the columns if necessary. An outside index would be necessary to designate proper reels. The module 12 finds the designated (by punching into another keyboard) accession number by reading the film at ten feet a second. Printouts may be obtained within seconds of all or part of the display on the screen. Cost for the complete system is $40,353 (this includes camera, encoder, verifier/display, processing and developing equipment, reader, lens, zoom, module 12, printer). However, this cost could be cut to approximately $15,800 if use could be made at night of the equipment in the Registrar's office.

Reasons rejected:

1. High equipment costs
2. Have to train personnel to expose and develop film on Miracode equipment
3. Film sensitive to temperature changes and physical damage; equipment also must be carefully monitored.
4. Articles on the same subject would be separated.
5. If Registrar's equipment were used, the morgue may be put into a dependent position which would be undesirable.

2. Magnetic tape storage

Articles would be transferred by means of a software program from disk storage (currently in use for editing purposes) to storage on magnetic tapes. Stories would be retrieved with the use of an on-line index which would search the stories on tape. This would require the use of the tape drive already available in the Texan computing facilities as well as someone to mount the tape when it was needed for searching. A software program would also be required for search procedure and use of the index. Cost (outside of the software programs needed) would be minimal since magnetic tapes cost only $9 each and only one tape would be able to store approximately thirteen months of articles.

Reasons rejected:

1. Excessive time needed to search a magnetic tape--this is not an efficient way to retrieve information.
2. Equipment not compatible for the projected heavy use: tape drives would not stand up to this kind of use.
3. Tape drives would warp tapes themselves with this much use.
4. Tapes subject to wear and environmental effects, such as dust.

3. Computer output microform (COM); CMR, Inc.--Sutton Page

Articles would be transferred by means of a software program from disk storage (currently in use for editing purposes) to storage on magnetic tapes. The tapes would then be processed out-of-house to produce computer output microform (film or fiche) for use in the morgue. Reader/printers would be available for paper copies of needed stories. Cost would be minimal because no equipment would be needed, magnetic

tapes cost only $9 each and the microform is inexpensive when produced in this way ($1.50/fiche (96 frames) at CMR, Inc.).

Reasons rejected:

1. Because of last minute changes, stories as stored on disk are not necessarily what actually appears in the newspaper.
2. Archival film would still need to be made because only individual stories would be processed to COM.

4. Questicon (ANCIRS)--Image Systems, Inc.

Microfiche would be produced COM and then stored in a terminal where it would be automatically retrieved with the help of an index, also on microfiche and stored in the terminal. The terminal itself interfaces with the computer and retrieves needed articles almost instantly after the accession number is punched in at the terminal. Terminals are about $ 10,000 each. The cost of the fiche and needed software is in addition to this. Hard copy (printed copies) may also be obtained from the terminals.

Reasons rejected:

1. For the low volume and use which could be projected for the Texan morgue (relatively speaking) this system is not cost-effective.
2. Articles on the same subject would be separated, requiring additional searches of the index and more keying of accession numbers.
3. The reduction ratios used for the fiche used in this system are such that the fiche could not be produced locally.

5. Use of the University's CDC 6600/6400

Because of the greater capacity available with the University's CDC 6600/6400, some thought was given by the study group to effecting an interface with that system before

implementing any of the systems proposed herein which use the computer.

Reasons rejected:

1. It is doubtful that an outside computer facility could provide the amount of on-line disk storage required for the automated systems or the amount of "up" time.
2. Charges for using the outside computer would probably exceed the cost of using the TSP computer.
3. If the outside computer facility changed hardware or software, the morgue computer programs could require extensive revision or even replacement.

APPENDIX E
GLOSSARY

authority list: thesaurus or other list of terms which may be used for indexing; also called "control list"

circulation: controlled use of clippings and/or other morgue materials by patrons (users) by a systematized means of charging materials in and out of the morgue

keyword index: an index compiled of terms taken directly from the material (in this case the story) by means of a predetermined formula (e.g. mathematical frequency curves or linguistic position of the word)

microform: general term used to describe film on reels, cards (fiche), or in jackets that bears a miniature photographic copy of a document which may be read by special reader.

user: for the purposes of this study the primary user has been defined as Texan editors, reporters, and writers; secondary users are all remaining journalism students, faculty, and staff; tertiary users include all other possible persons whose information needs may be fulfilled by means of the Texan morgue

APPENDIX F

Following is a sample of each of the computer-generated printed indexes which have been recommended in the preceding sections of the study. The first is the index which would be used by the Texan staff itself, the subject heading index. The second example is that of the index which would be the complete index, containing all of the information of the subject heading index as well as 50-character headlines of the articles.

R-T-F, Department 2345
02037645678
02037645689
02127623456

Regents, Board of 2346
03047523456
03247534567

Rogers, Lorene 2347
10056923456
11127024354
01277186745

Swahili Club 2348
12087423456
05087554657

Texas Student Publications 239
02037645678
02127623456

R-T-F, Department 2345
02037645678 Changes in RTF Curriculum
02037645689 RTF Welcomes faculty members

Regents, Board of 2346
03047523456 Governor Appoints Three Regents
03247534567 Students Portest Regents' Decision

Rogers, Lorene 2347
10056923456 Rogers Meets with Texan Representatives
11127024354 Texan Awaits Roger's Decision

AFTERWORD

While the study group as a whole is responsible for the conclusions and recommendations made, certain members have greater depth of background in the areas listed below. They may be contacted for further information at the numbers listed or via messages left with Professor E. B. Jackson, Graduate School of Library Science, Centrex 3821, PAX 2142.

Manual systems	Amanda Masterson Trent Schmitt	478-1121 478-2060
Automation by computer	Tom Kirwin	442-0908
Costs	Sulinda Cole Gary Kansteiner	459-8961 476-8169
Implementation	Miguel Bamberger Paul Strohm	453-0318
Microforms	Mickie Voges*	459-8894

The study group trusts that this example of a class project done in one School for an intended useful purpose in another area of the University community is not unique. We believe that the benefits derived from this practice are to both parties; we would like to acknowledge again the help and cooperation received from TSP and encourage this type of interchange within the University community in the future.

* Final editing of the text and reproduction master preparation was by Mickie Voges